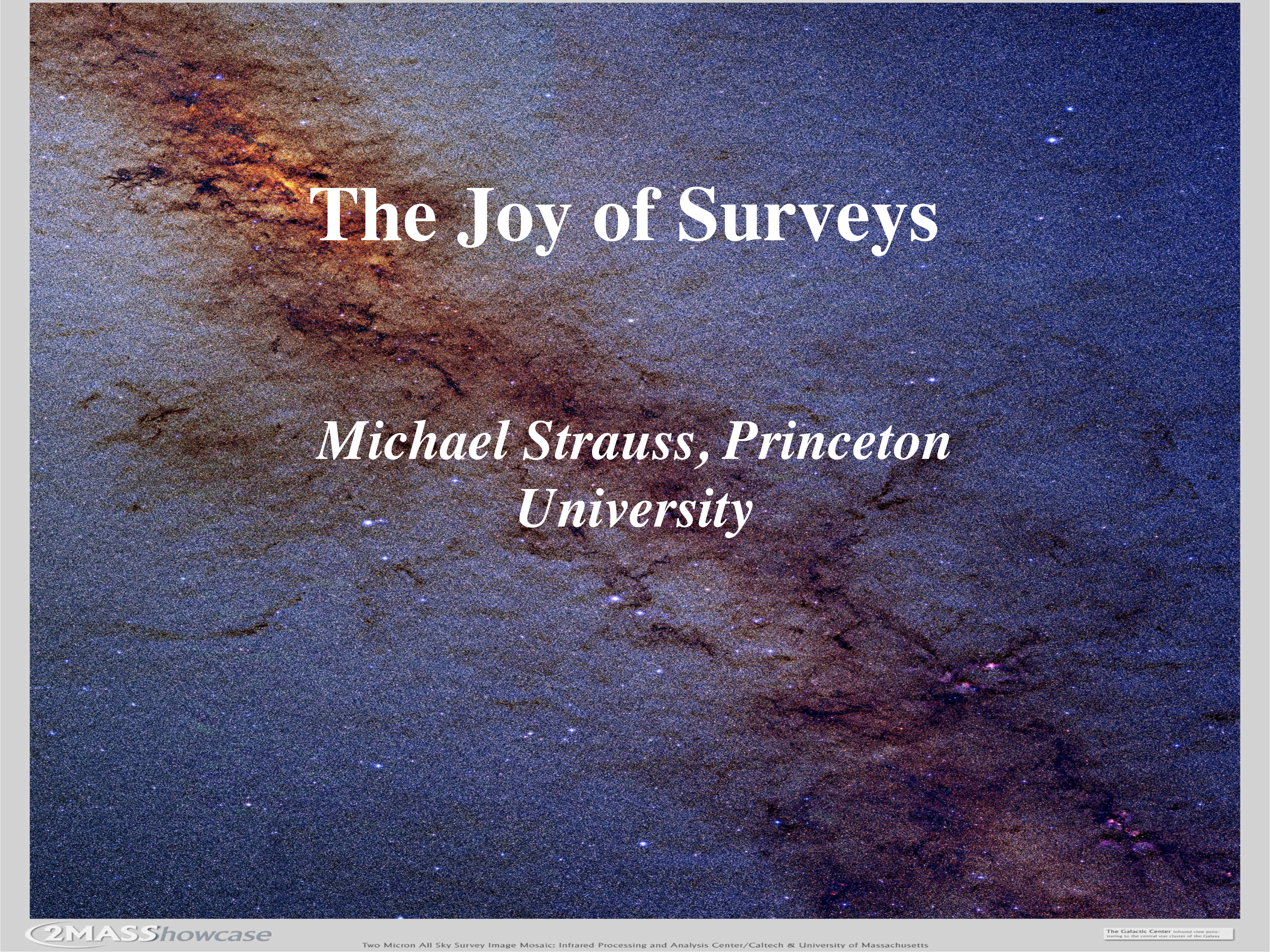
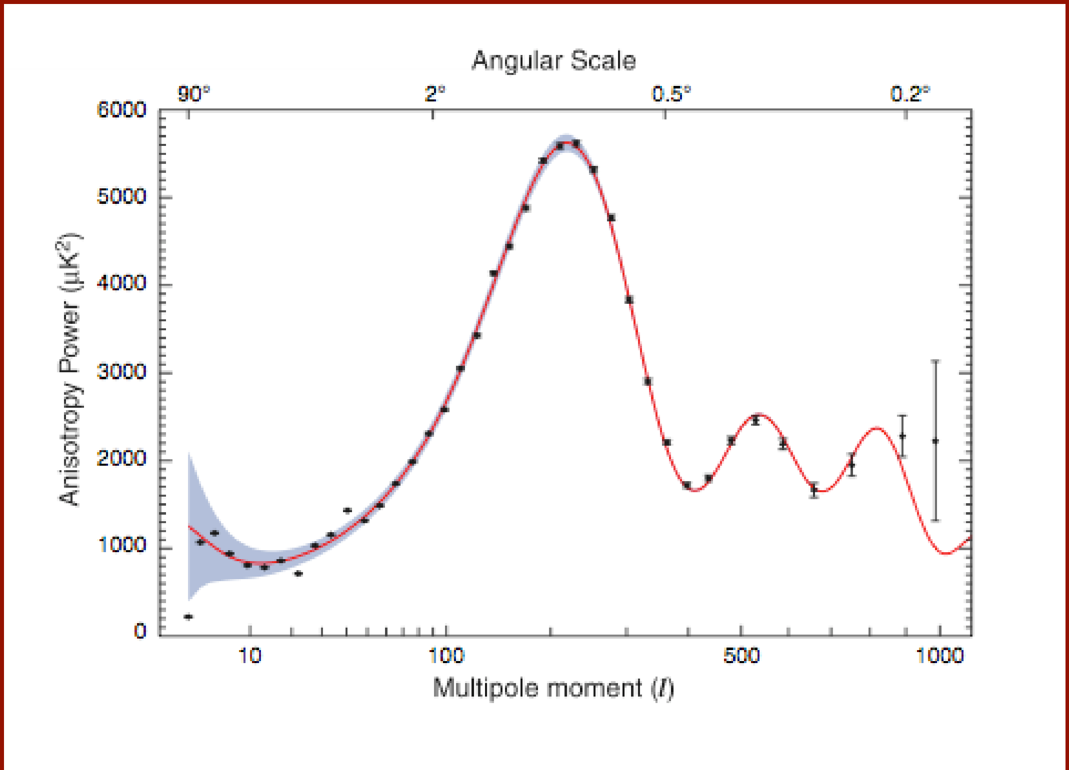
The Joy of Surveys

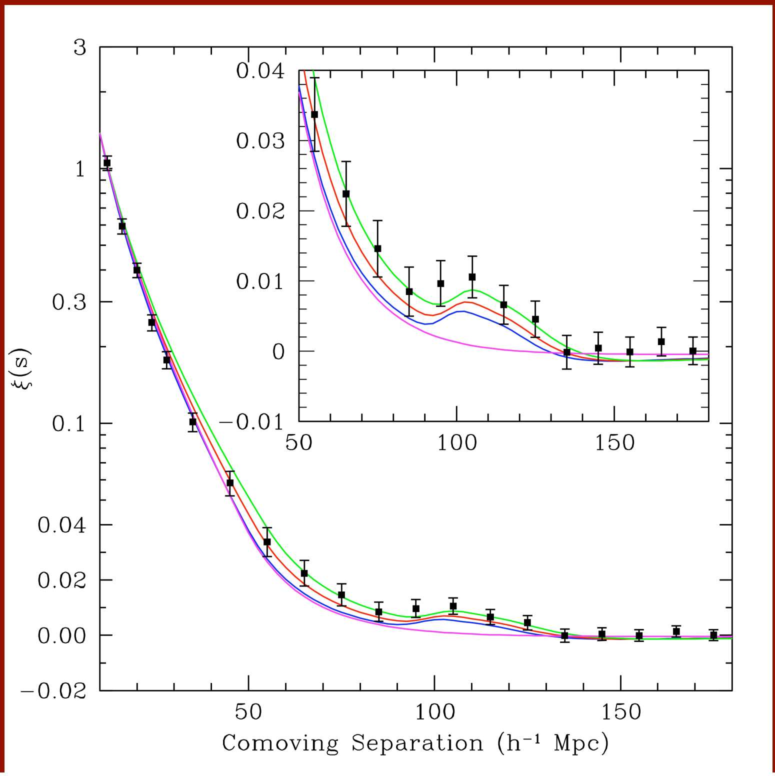
Michael Strauss, Princeton University



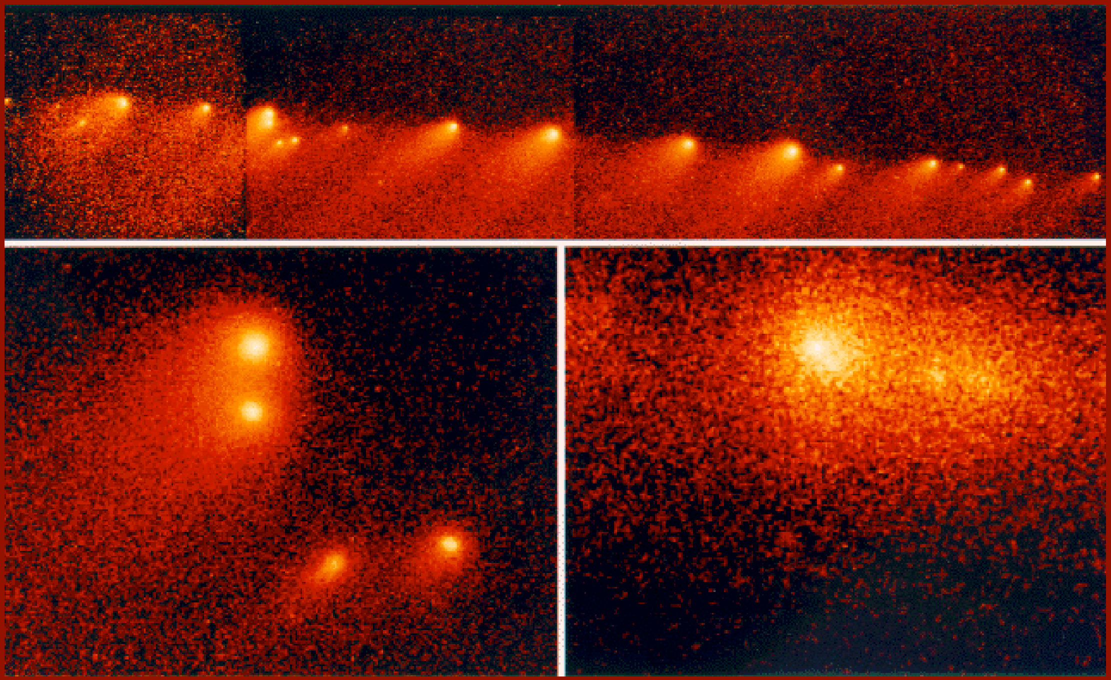
Why do we do astronomical surveys?

- To address very specific scientific goals, such as to measure baryon oscillations in the galaxy distribution as a probe of dark energy (talks by ***Eisenstein, Seljak***)



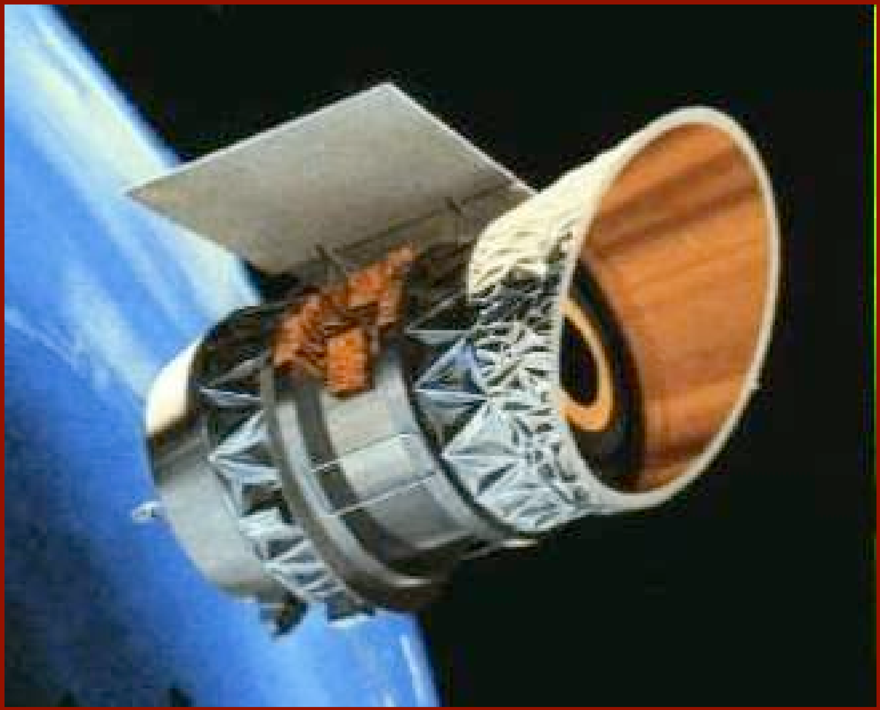


or to search for killer asteroids



▪ To take advantage of new technological capabilities

- IRAS



▪ GLAST/Fermi

Science goals and new technology are tightly coupled. Designing a new survey:

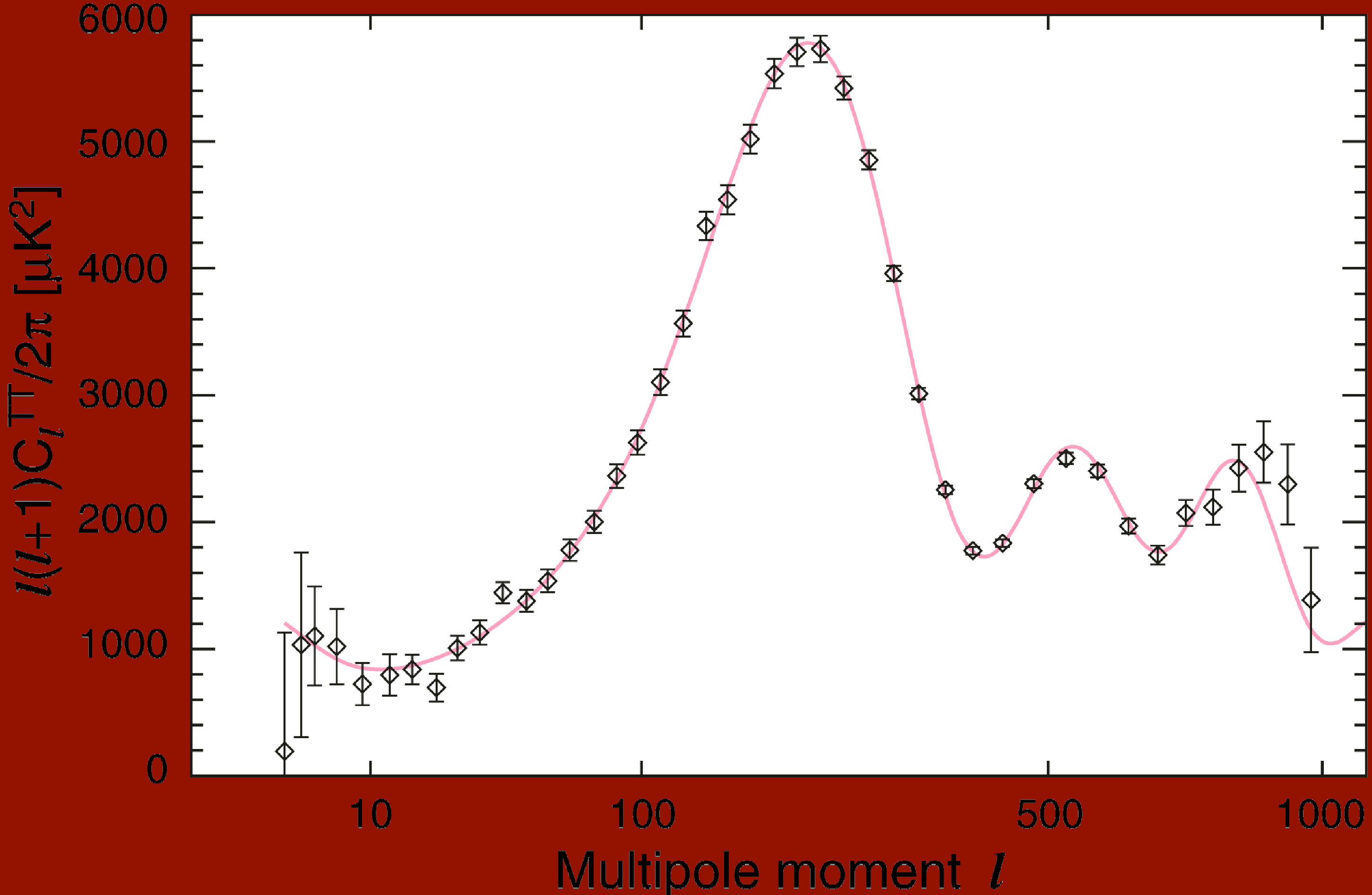
- Figure out what the basic science goal is:
  - Survey depth?
  - Error bars on wide-angle power spectrum?
  - Completeness of near-Earth asteroid sample?
- Decide on parameters of telescope, instrument, survey cadence and software.

- Given your hardware complement, determine the *physical* limits on accuracy.
- Work like hell (but no harder) to approach these limits.
- Doing things right the first time saves enormous trouble down the road!

Quality, Quality, Quality

- Unless you push to the physical limits of your data, you will never be able to ferret out subtle systematics, and you will give up potentially valuable science goals. (*It is statements like this that drive project managers crazy*…)
- Doing so will enable a large range of science well beyond the original science goals.

WMAP finds χ^2/dof ~ 1.06



Nolta et al. 2008

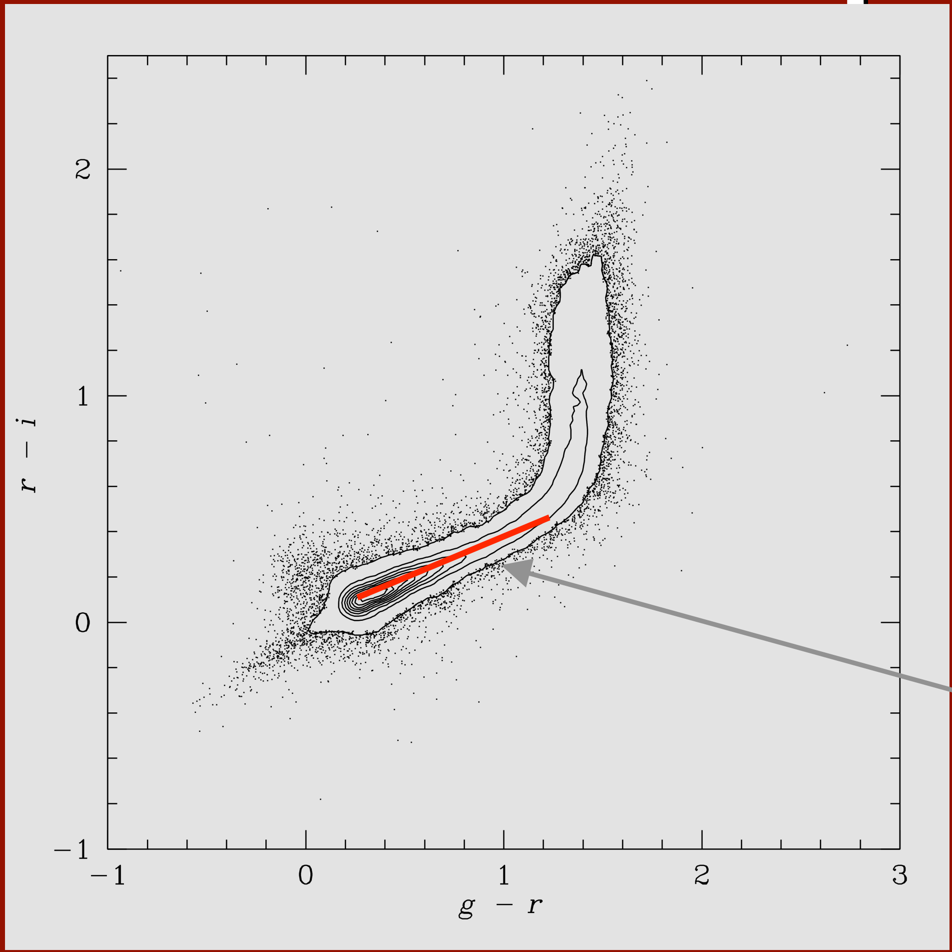
Case in Point: SDSS

- Original (advertised) motivation: A million redshifts to measure the large-scale distribution of galaxies.
- Requirement to select galaxies in a uniform way required:
  - Imaging several magnitudes deeper than spectroscopic limit;
  - Highly uniform photometric calibration;
  - High success rate in redshift determination.
- A happy coincidence was that a 2.5m telescope worked well for both the imaging and spectroscopic components of the survey.

Pushing to the limits of the data enables unanticipated science

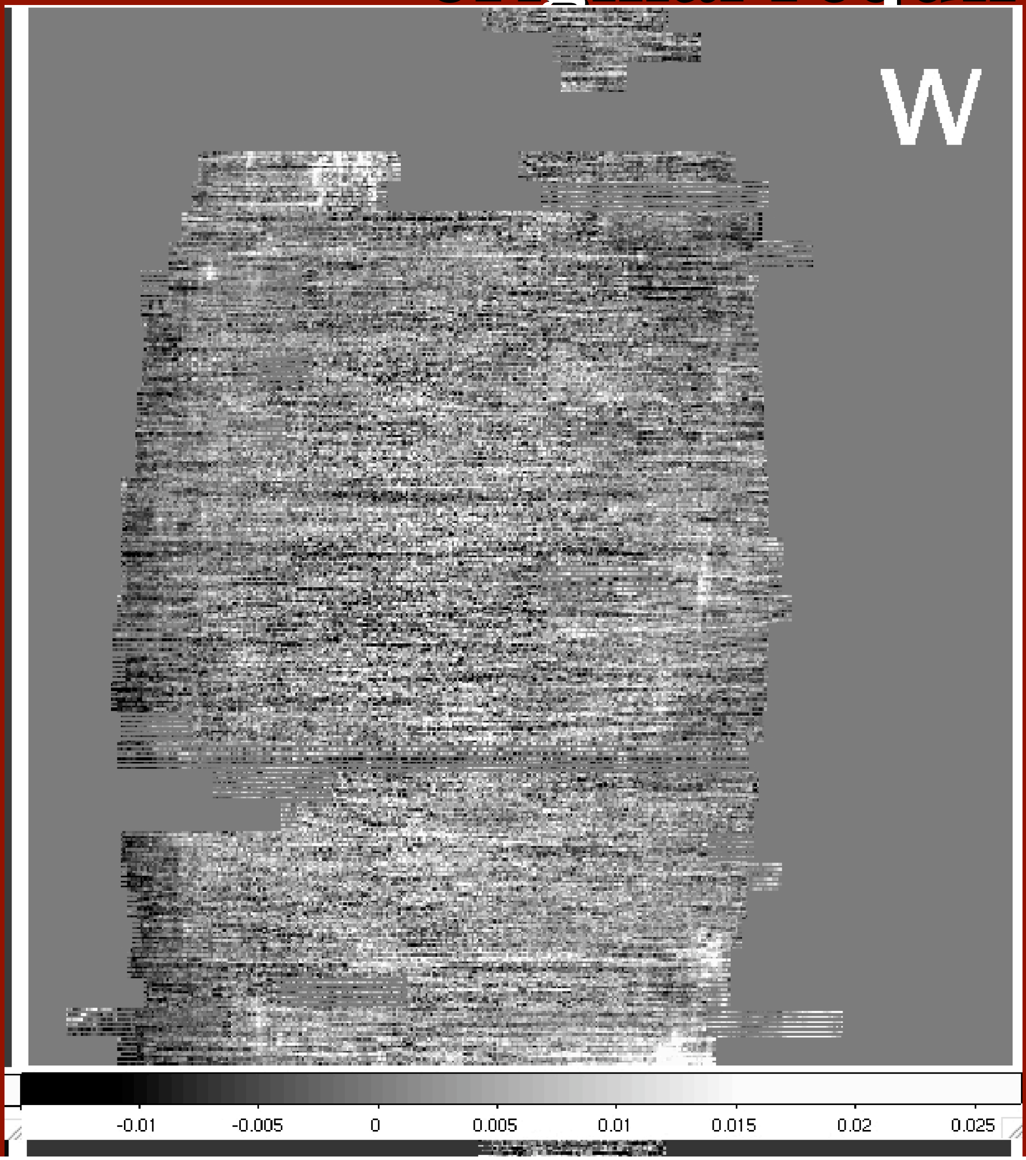
- For example: SDSS set a photometric uniformity goal of 2% in *g*, *r*, and *i*, and 3% in *u* and *z*.
- This turned out to be difficult for all sorts of reasons, including:
  - Getting the flat-field right
  - Getting the point-spread function right
  - Getting the standard stars right
  - Understanding the filters
  - Understanding the atmosphere
- A many-year effort sorting all of this out!

Distribution of stars in color-color space



The position of the ridgeline of the stellar locus is a powerful test of the uniformity of photometric calibration

The position of the stellar locus has rms variations of ~1%, twice as good as original requirements!

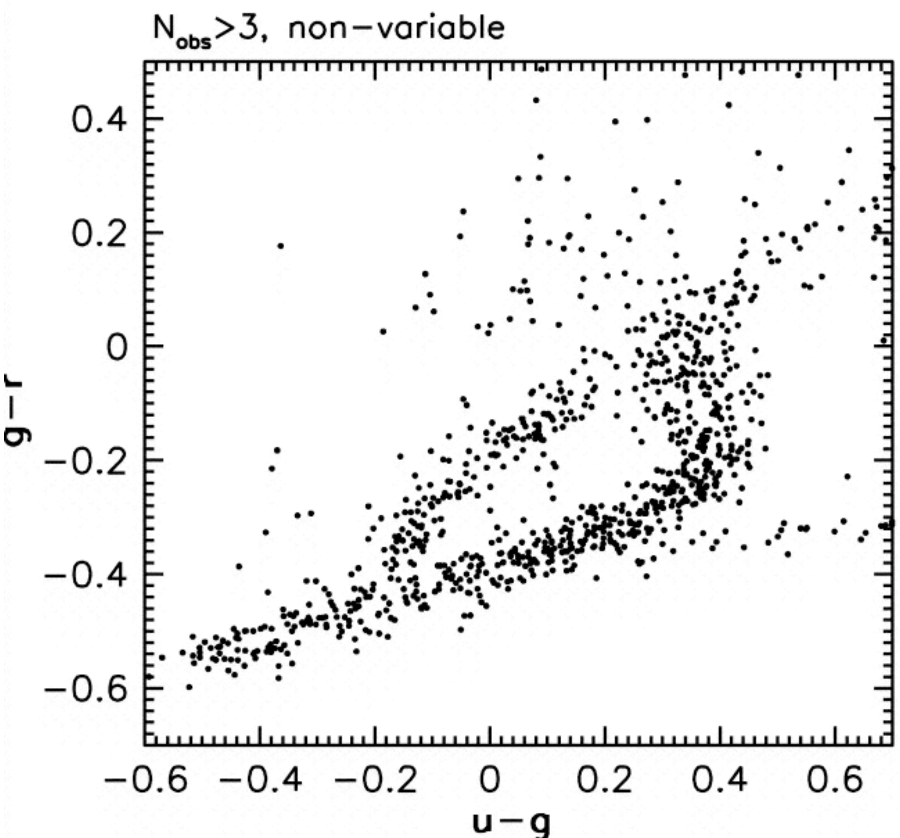


Variation in position of stellar locus across survey area. Plotted scale ranges ± 2% (0.02 mags). Photometric calibration is done by forcing agreement in overlaps between runs (similar in spirit to CMB experiments like WMAP)

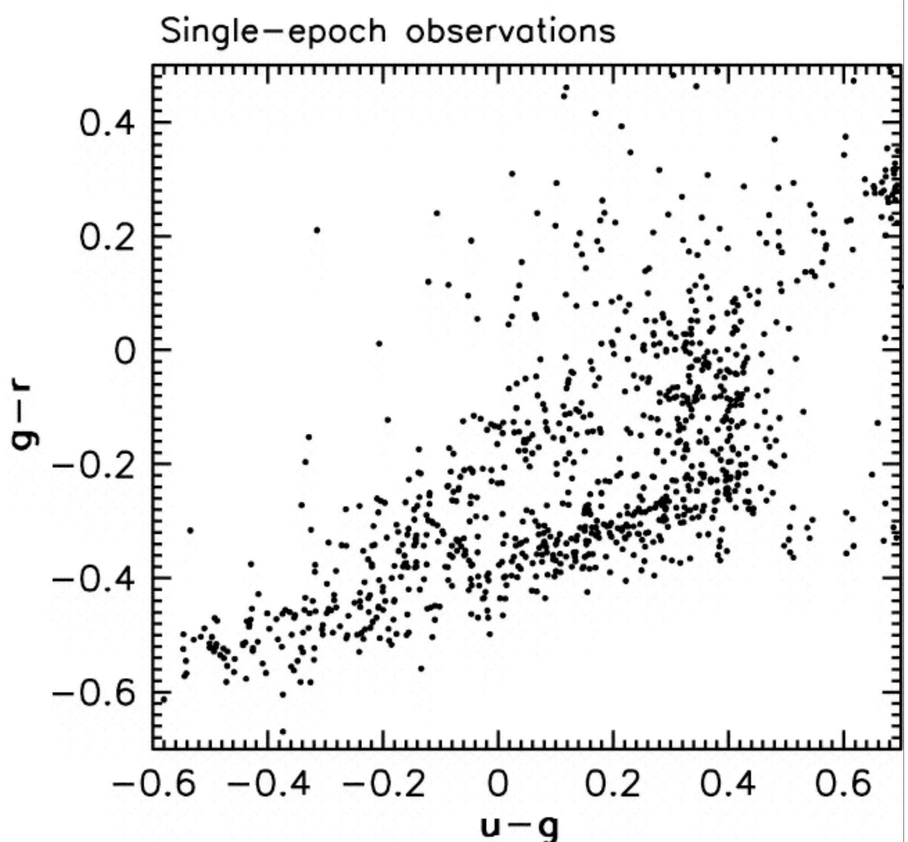
Padmanabhan, Finkbeiner, Schlegel et al. 2008

Color-color diagram of white dwarfs

1% photometry



2% photometry



Getting the data out?
Getting the data right?

- A balance: you want to get the data out to your collaborators, and to the world, as soon as possible. The best way to find the flaws in your data is by doing science with it. But you don't want to be distributing data that you know are overly flawed…
- And don't underestimate how long commissioning will take, especially when you have new hardware.

A lesson: Don’t hold your data too closely to your chest

- For major surveys, there is typically much more science to be done than your collaboration can carry out. Don’t obsess about getting scooped!
- Major surveys are expensive enough that you need the resulting good will from the community to get them funded.

- The act of getting the data into shape to distribute publically makes it much easier for collaboration members to get to the data as well. The era of flat files is over; we all have to become data miners. See presentations by *Szalay*, *Connolly*.
- Bottom line: SDSS would have been much less successful if we hadn't been forced to make the data public!

The Importance of Software

- Data pipelines are an absolutely core component of any survey. They need to be planned for from Day 1, and are often one of the most technically challenging parts of he project. A professional approach! In the beginning, SDSS imagined that the pipelines would be written by faculty members in their spare time… See talk by *Lupton*.
- Plans for databases and data distribution (both within the collaboration, and to the world) also need to be built into the survey design.

Big Surveys are Difficult Politically

- Large groups of people have to work together for years. How do they all get along? How will those working on infrastructure get rewarded professionally? Everyone's favorite question: *Who gets to be first author?*

Major Themes for the Next Generation of Surveys

- **The Time Domain**. The heavens are not static! Variables, transients, proper motions, planetary transits, parallaxes, asteroids, oh my… An increasingly important theme at all wavebands.
- We need resources for follow-up. Think GRB afterglows, which require observations from gamma-rays to radio. See talks by *Kulkarni*, *Wozniak*, *Djorgovski*, *Bower*, and *York*.

Cosmology

- Baryon oscillations, weak lensing, supernovae, CMB fluctuations, Reionization, etc. etc. See talks by ***Seljak, Lawrence, Benson, Padmanabhan, Jain, Furlanetto, Heitmann, Eisenstein, Refregier, Backer, Malhotra, Church***… A major theme of this meeting!

Galaxy Evolution

- Watching the mass build-up in galaxies, understanding how they got their morphologies. Talks by *Gawiser, Carilli, Yun, Mohr, Windhorst, Henning*.
- **Galactic Structure and Stellar Populations:** under-represented in this meeting.
- **Planets and asteroids:** Talk by *Heasley*

Synergy between surveys

- The ability to study the physical nature of objects across very large wavelength ranges is becoming ever more important. For example, the GOODS and AEGIS surveys included observations from X-ray to radio and everything in between… Encouraging interaction, coordination, and inter-operability between surveys will be a major theme of this meeting.

Follow-up!

- Rare objects (high-z quasars, unusual stars, etc).
- Spectroscopy! In the optical, there are many major imaging surveys planned, culminating in LSST. Next-generation spectroscopic surveys are perhaps not quite as ambitious (but see SDSS-III, LAMOST, HETDEX, ADEPT, WFMOS?). Let's have some discussion about the direction major spectroscopic surveys should take.